

HTCondor in the enterprise

Ben Cotton

Senior Support Engineer





Photo by Dmitry Avdeev. Used under CC BY-SA 3.0 Unported



We believe utility access to technical
computing power
accelerates discovery & invention



About us

- We like to say: Better Answers. Faster.
- We want our customers to get the resources they need when they need them



You may recall...



CycleComputing.pptx

150%

Search in Presentation

Home Themes Tables Charts SmartArt Transitions Animations Slide Show Review

Slides Layout New Slide Section Font Paragraph Insert Text Picture Shape Media Arrange Quick Styles Slide Show Play

Slides Outline

25 We did it!
• Went from 0 to 50k cores in 23 minutes
• Peaked at ~70K cores from 5689 instances
• Simulation completed in 8 hours
• Infrastructure cost: \$5,554

26 Where do we go from here?

27 Better-er answers. Faster-er.

28 If you build it, they will come
• Large financial institution actuarial modeling
– Originally just wanted to do Federal Reserve stress tests
– Then mid-sized actuarial firm
– Now regularly use 8000 cores in AWS

29 Coming concerns
• Data movement
• Multi-provider cloud usage
• Seamless burst to cloud

30 We write software to do this...

Click to add notes

Normal View Slide 27 of 32 150%

Large life insurance company



In the old days...

- Departments decided to pool their resources
- 10x the capacity, but now they have to *share*
 - Group quotas and preemption



In the old days...

- It was all worth it
 - Month-end reports finish in less than a month
 - The “scheduler” is no longer a senior actuary with a clipboard



Then FDIC requirements happened

- Needed to run a business-critical workload in a defined period of time
- No time or space to add to internal capacity
- 8000 AWS cores for 2 weeks
 - P.S. they passed the stress test



This cloud thing is sort of cool...

- 8000 AWS cores for month-end runs
 - Much shorter time-to-results
- We built it, they came



Always be computing

- Usage moved cloudward
- Internal datacenter was decommissioned
- Added another pool for development work
- There's no “month-end” anymore
 - All the analysis, every night



What they like about HTCondor

- Reliable and fault-tolerant
- Well-suited for batch
- Well-suited for cloud
- Multi-stage workflows
- Can implement business policies



Large hard drive manufacturer



You may remember this one

- Moved from 256 core in-house cluster to 8000 core AWS cluster
 - Brief stopover at a 78k core cluster
- Parameter sweeps to model hard drives



You may remember this one

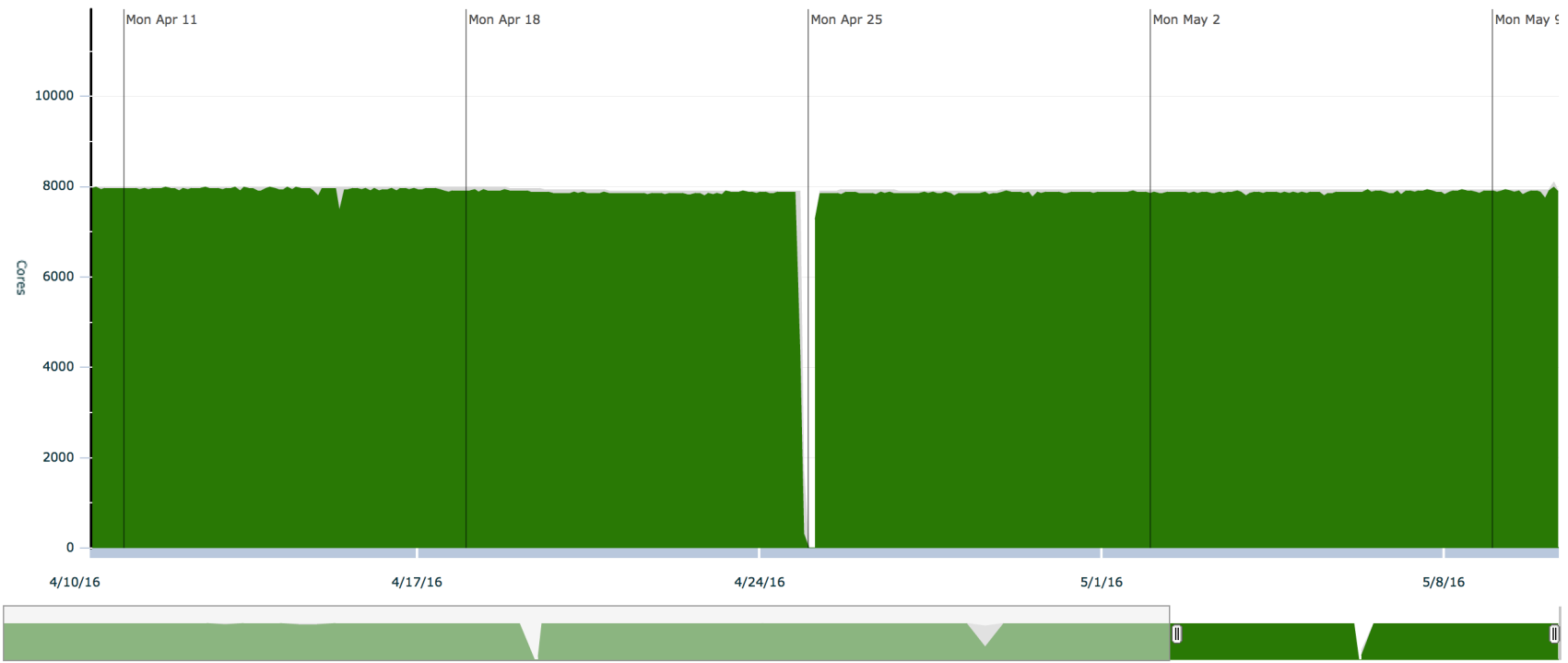
- Moved from 256 core in-house cluster to 8000 core AWS cluster
 - Brief stopover at a 78k core cluster
- Parameter sweeps to model hard drives





Time Frame: 3 Hours | Day | Week | Month

View as: Area | Line



Legend

cyclecloud

Unclaimed



Getting more work done

- Modeling changes faster to reduce lost expenses
- Adding more users



What they like about HTCondor

- Reliable and fault-tolerant
- Well-suited for batch
- Well-suited for cloud
- Easy to move jobs to the front of the line



Life sciences SaaS platform



DNA and RNA sequencing

- Small cluster doing important work
- Future plans
 - Heterogeneous resources
 - DAG-based workflows



What they like about HTCondor

- Well-suited for batch
- Well-suited for cloud
- Support for heterogeneous resources
- Multi-stage workflows



I'm sensing a theme...



Well-suited for batch

- High-throughput computing is used for business-critical workloads across many industries
- (But, HPC workloads are important, too)



Well-suited for cloud

- Resources can disappear with little warning
- Moving from internal to cloud isn't a complete change in the technology



Multi-stage workflows

- Automate as much as you can
 - But no more



Business rules rule!

- Fair-share is almost never what you want
- Abusing quotas for fun and profit (but mostly profit)



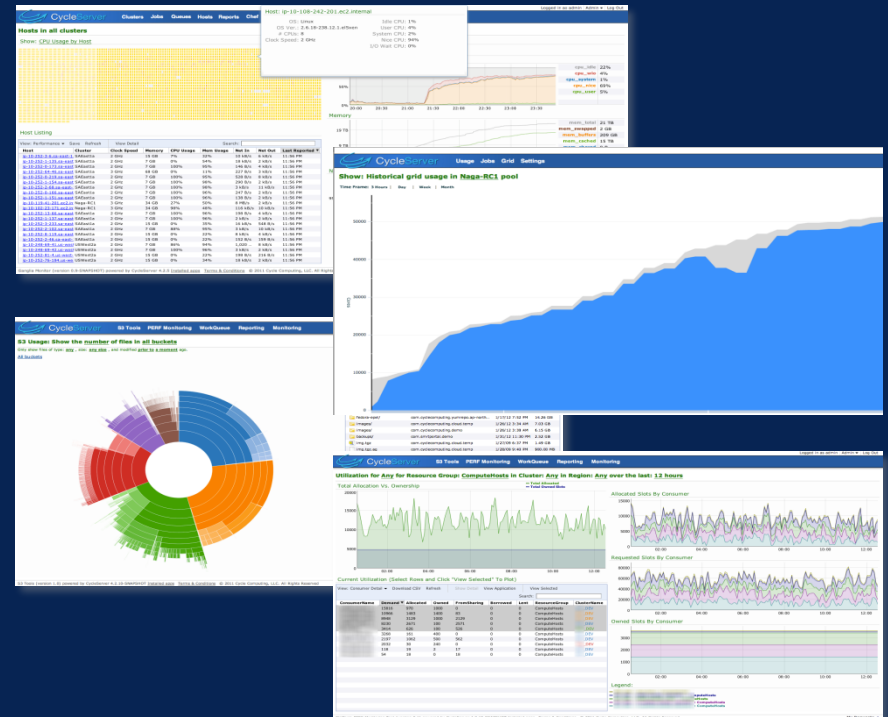
What's next?



We write software to do this...

Cycle Computing easily orchestrates workloads and data access to local and Cloud technical computing

- Scales from 100 - 100,000's of cores
- Handles errors, reliability
- Schedules data movement
- Secures, encrypts and audits
- Provides reporting and chargeback
- Automates spot bidding
- Supports Enterprise operations



Does this resonate with you?



We're hiring support
engineers, solutions
architects, sales, etc.

jobs@
cyclecomputing.com



Reception tonight!

- Free tacos! Free beer!
- 6-7 PM at Lucky's Bar & Grille
 - 1421 Regent Street

